

CLAIMS

What is claimed is:

1. A method of reducing a number of write operations relative to delivery of out-of-order RDMA Send messages, the method comprising the steps of:

providing a reference counter with a completion queue element (CQE);

setting the reference counter to a number of RDMA Send messages completed for a selected TCP hole;

reducing the reference counter by one for each Poll-for-Completion conducted by an RDMA verb interface; and

removing the CQE from a respective completion queue (CQ) where the counter becomes zero.

2. The method of claim 1, further comprising the step of updating a work queue element (WQE) of a send queue only if it has greater than a threshold number of pending RDMA Send messages.

3. The method of claim 2, wherein the threshold is commensurate with resources allocated for storing information for pending RDMA Send messages.

4. The method of claim 1, further comprising the step of having the CQE include at least part of completion data.

5. The method of claim 4, wherein a remainder of the completion data is included in a work queue element (WQE) associated with one or more RDMA Send messages.
6. The method of claim 4, further comprising the step of indicating one of: 1) the CQE includes all completion data, and 2) CQE completion data is at least partially included in a work queue element (WQE) associated with one or more RDMA Send messages.
7. The method of claim 1, wherein a number of write operations equals $N+1$, where N is a number of completed RDMA Send messages that were pending prior to the selected TCP hole closing.

8. A system for reducing a number of write operations relative to delivery of out-of-order RDMA Send messages, the system comprising:

means for setting a reference counter of a completion queue element (CQ) to a number of RDMA Send messages completed for a selected TCP hole;

means for reducing the reference counter by one for each Poll-for-Completion conducted by an RDMA verb interface; and

means for removing the CQE from a respective completion queue (CQ) in the case that the counter becomes zero.

9. The system of claim 8, further comprising means for updating a work queue element (WQE) of a send queue only if it has greater than a threshold number of pending RDMA Send messages.

10. The system of claim 9, wherein the threshold is commensurate with resources allocated for storing information for pending RDMA Send messages.

11. The system of claim 8, further comprising means for storing at least part of completion data in the CQE and a remainder of the completion data in a work queue element (WQE) associated with one or more RDMA Send messages.

12. The system of claim 11, further comprising means for indicating one of: 1) the CQE includes all completion data, and 2) CQE completion data is at least partially included in a work queue element (WQE) associated with one or more RDMA Send messages.

13. The system of claim 8, wherein a number of write operations equals $N+1$, where N is a number of completed RDMA Send messages that were pending prior to the selected TCP hole closing.

14. A computer program product comprising a computer useable medium having computer readable program code embodied therein for reducing a number of write operations relative to delivery of out-of-order RDMA Send messages, the program product comprising:

program code configured to set a reference counter associated with a completion queue element (CQE) to a number of RDMA Send messages completed for a selected TCP hole;

program code configured to reduce the reference counter by one for each Poll-for-Completion conducted by an RDMA verb interface; and

program code configured to remove the CQE from a respective completion queue (CQ) in the case that the counter becomes zero.

15. The program product of claim 14, further comprising program code configured to update a work queue element (WQE) of a send queue only if it has greater than a threshold number of pending RDMA Send messages.

16. The program product of claim 14, wherein the threshold is commensurate with resources allocated for storing information for pending RDMA Send messages.

17. The program product of claim 14, further comprising program code configured to store at least part of completion data in the CQE.

18. The program product of claim 17, wherein the storing program code stores a remainder of the completion data in a work queue element (WQE) associated with one or more RDMA Send messages.

19. The program product of claim 17, further comprising an indicator of one of: 1) the CQE includes all completion data, and 2) CQE completion data is at least partially included in a work queue element (WQE) associated with one or more RDMA Send messages.

20. The program product of claim 17, wherein a number of write operations equals $N+1$, where N is a number of completed RDMA Send messages that were pending prior to the selected TCP hole closing.